



THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 1997/18

Learning with a Known Average:
A Simulation Study of Alternative Learning Rules

by

Huw D Dixon and Paolo Lup

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Learning with a known average: a simulation study of alternative learning rules.

Huw D. Dixon*	Paolo Lupi
Tel: (44) 1904 433788	Tel: (44) 1904 434676
email: hdd1@york.ac.uk	email: pl105@york.ac.uk

Economics Department,
University of York,
York YO1 5DD
UK
December 1996

Abstract

In this paper we consider an economy consisting of identical homogeneous good Cournot duopolies. Firms decide whether to experiment or not on the basis of their aspiration level. If they decide to experiment, then there is a switching rule which determines the output they will produce next period. The switching rules that we consider are: (i) random switching, (ii) imitation, (iii) best-response dynamic. The aspiration level is given by the average level of profit in the economy, as in Dixon (1996), possibly subject to some noise. We find that the random switching and imitation leads to all duopolies becoming joint profit maximizing, even with a small level of noise. Best-response dynamic leads to all duopolies becoming Cournot even with very small levels of noise. With high levels of noise the different learning rules lead to different and diverse limiting populations. We are able to compare the different limiting populations and the dynamic path.

*This is a preliminary draft: comments are welcome. We would like to thank conference participants at the II Workshop on Economics with Heterogeneous Interacting Agents held at Ancona, Italy (May '97) and at III International Conference on Computing in Economics held at Stanford CA (July '97). This paper is part of an ESRC funded project "Evolution, Oligopoly and Competitiveness" (grant R000 23 6179).

1 Introduction

In this paper, we explore the framework of Dixon (1996) in the context of Cournot Oligopoly. Consider an economy with many markets, each with the same cost and demand conditions. In each market there are two firms playing a Cournot duopoly. There is a capital market in the economy, or some institution which measures the average level of profit in the economy as a whole. This average can be thought of as the “normal” level of profits in the economy. The firms are modeled as being boundedly rational, and we adopt an aspiration-based model of decision making by the firm. The firms have an aspiration level: if they are attaining their aspiration level, then they are unlikely to change their action; if they are below the aspiration level, they are very likely to change their action. The aspiration level is endogenous, and for each firm it is equal to the level of average profits. A dynamic system is thus set up: the current actions of firms determine their own and average profits.

Dixon (1996) analyzed this model from a theoretical and analytical perspective. Under two key assumptions, he was able to show that there is a unique and nearly global attractor to the above system: the collusive or joint profit maximizing (JPM) outcome. This analytical result is in some ways very general, but also has limitations. The dynamic process is represented as a non-stationary Markov chain, and these are notoriously difficult to analyze. The purpose of the current paper is to adopt more general assumptions than those made in the previous paper, and to evaluate the model using simulation techniques.

There are two stages to the firms decision process at each period in time. First, there is the decision whether or not to experiment. In this paper, “to experiment” means to (attempt) to change action¹. In Dixon (1996), this was deterministic: if you are above the aspiration level, you do not experiment; if you are below you do. We introduce “noise” into this process. In particular, we have δ -noise, where δ is the probability that you experiment if you are above average, and also β -noise, where β is the probability that you do not experiment when you are below the aspiration level. Aspiration based models of learning are not new (see, inter alia² Lewin (1936), Simon (1947, 1981), Siegel (1957), Borgers and Sarin (1994), Palomino and Vega-Redondo (1996)).

Second, given that you have decided to experiment, there are the conditional switching probabilities. These give the probability of switching to

¹We say attempt, because you may by chance end up choosing the same action.

²See Kornai (1971) for a discussion.

the set of available actions, conditional on deciding to experiment. In Dixon (1996) these were analyzed under fairly general assumptions, but without β or δ noise. In this paper we analyze some specific switching rules: random switching, imitation, best response and their mixture. Random switching (RS) means that if a firm decides to experiment, it chooses each strategy with the same probability (including its current strategy). Imitation (IM) means that the firm randomly looks at another firm, and copies it (all firms are equally likely to be observed as in chapters 4, 5 of Weibull (1995) and in Schlag (1996b)). Both imitation and random switching are strategically “blind”: the firm does not look at what its competitor is doing and work out a sensible strategy to follow. Best response dynamics (BR), however, means that the firm chooses its best response to the action played by its opponent in the previous period as in the classic Cournot model of adjustment.

The closest theoretical models to the one explored in this paper are Bendor and Mookherjee (1994), Kandarikar et al. (1995) and Palomino and Vega-Redondo (1996). The first two papers require that the aspiration level is constant and equal to the long-run population average payoff. Palomino and Vega-Redondo has the aspiration level equal to the current average payoff. However, their model has a random matching technology and looks specifically at the two strategy prisoners dilemma (PD). All of these papers find that the cooperative outcome is possible as an equilibrium: in Dixon (1996), however, we find that under certain assumptions the cooperative outcome is the *only* possible outcome.

The advantages of the simulation methodology in understanding this model are several. Firstly, the analytical results in Dixon (1996) which underlie the paper are concerned with the asymptotic properties of the dynamic process. Simulation enables us to understand the dynamics of adjustment, both in terms of the speed and the profile. Secondly, δ and β noise are crucial in determining the asymptotic and general behaviour of the system. Since we currently have no analytical results for this, the simulations are able to tell us if the no-noise analytical results are robust to having small noise levels, and also how the system behaves when the noise increases. Thirdly, the effects of the “initial position” can be explored in some detail.

We will give a full summary of the results in section 3 of the paper. However, the general conclusions can be summarised as follows:

1. The analytical results are robust to small amounts of noise for the three switching rules we consider. Both imitation and random switching give rise to the JPM (joint profit maximizing) outcome: best response dynamics gives rise to the Cournot outcome.
2. With high levels of noise, the three different switching rules give rise to

different dynamic properties. (a) With RS, the end point distribution contains many surviving firm types, and is not very sensitive to initial position. (b) With IM, the end distribution tends to involve firms in most markets choosing the same output level, and the economy is often more competitive than the JPM outcome. However, with IM the initial position matters to a great extent when there is a high noise level. (c) With BR, we find that for low levels of noise the Cournot outcome is the most common outcome. Unlike the other two cases, however, we find that asymmetric market structures are possible with BR: surviving market structures can involve firms producing different levels of output.

Whilst these results have been derived for a specific model — Cournot Duopoly — we believe that they are not atypical. In a companion paper Lupi (1997) we conduct a similar analysis for simpler games with only a few strategies (prisoner’s dilemma, coordination games), and obtain analogous results.

These results are, we believe, important to economists. Certain markets, most notably the financial markets, convey information (implicitly or explicitly) about the whole economy. In particular, they convey information about the average or normal rate of return on capital. This information will effect the way that individual firms behave. Whilst the model we have developed is very simple, it does show how this mechanism can give rise to a social learning process. The learning process is social in that its driving force is the population mean payoff: firms who perform below average are forced to do something (by shareholders, or other of the various mechanisms constraining managers).

2 The Model

We have built a general framework in which to analyse the effect on the evolution of the population of three different learning rules: random switching, imitation and best response. We have also addressed the question of how the evolutionary process is affected by the introduction of “noise”.

Players, Strategies and Payoffs. We consider a symmetric two-player game with a finite set $S = \{1, 2, \dots, K\}$ of pure strategies or firm-types. If firm-type i plays firm-type j , then the payoff of firm i is $\pi_{i,j}$. Π is the $K \times K$ matrix of payoffs $\pi_{i,j}$.

\mathcal{L} is the set of all possible pairs of strategies, one for i and one for j . Since in our model it makes no difference which firm is type i and which is type j ,

we denote a pair of strategies by (i, j) with the convention that $i \leq j$. Hence the number of elements in the set \mathcal{L} is:

$$L = \sum_{i=1}^K i = K(1 + K)/2.$$

The one to one mapping $\ell(i, j)$ gives each pair (i, j) a unique number between 1 and L . We assume the form:

$$q = \ell(i, j) = j + (k - \frac{i}{2})(i - 1)$$

Hence we can write the set of pairs as $\mathcal{L} = \{1, 2, \dots, L\}$ with generic element q, r or $\ell(i, j)$.

We can think at every pair of strategies as at a location or market where a duopoly game is played between two competitors which use, respectively, strategy i and strategy j .

We consider discrete time $t = 0 \dots T$. In Dixon (1996) we consider $T = \infty$ in an analytical framework. In the current paper, T is the endpoint of the simulation which is in some cases endogenous (there is a stopping rule), and in others fixed ex ante.

There is a “population” of duopolies represented by the unit interval. The set \mathcal{L} can be used to partition the population of duopolies since each duopoly $d \in [0, 1]$ belongs to one and only one of the L pairs according to the strategies used by the two firms. We define the index function $I(d, (i, j), t)$ which takes the value 1 iff the firms at d play strategies (i, j) at time t . The proportion of duopolies which adopt a certain pair of strategies at time t is defined as $P_t(i, j)$, where:

$$P_t(i, j) = \int_0^1 I(w, (i, j), t) dw.$$

The proportion of firms of type i at time t , $\hat{P}_t(i)$ is:

$$\hat{P}_t(i) = \sum_{j \geq i} P_t(i, j) + \sum_{j < i} P_t(j, i).$$

The average profits in the “economy” are:

$$\bar{\pi}_t = \sum_{i=1}^K \sum_{j \geq i}^K P_t(i, j) [\pi_{i,j} + \pi_{j,i}] / 2.$$

We assume that each individual firm knows this economy wide average: this can be via capital markets, or a general notion of “normal” profitability.

Clearly, as the proportions $P_t(i, j)$ change over time, the average profits will vary.

Aspirations and Learning. In our model each firm adheres to a very simple behavioural rule that tells the firm what action to take at every time t . Each firm has an aspiration level α_t : if it is earning at least α_t , then it continues with its existing strategy with probability $1 - \delta$ (where $\delta \in [0, 1]$); if it is earning less than α_t , then it *experiments* with probability $1 - \beta$, by choosing a (possibly different) strategy. The parameters δ and β are noise parameters.

In the case of no noise ($\delta = \beta = 0$) the model is particularly simple. Each firm compares its own performance with its aspiration level α_t and uses this information to review its own behaviour. If a firm is performing below its aspiration level, it switches to another firm-type (that could be itself again) by means of a particular “switching technology”: if it is earning at least α_t , then it does not experiment.

The aspiration level in our framework is endogenous, and depends on average profits: $\alpha_t = \bar{\pi}_t$ ³. Since average profits vary over time with the evolution of the proportions, there will be an induced non-stationary dynamics of the aspiration levels of firms. This non-stationarity is the effect of the complex interaction between firms: firms are in fact observing each other either directly (like in the imitation case) or indirectly (via the average profits) and change their behaviour as they learn. This captures the idea, present in psychological literature that agents revise their aspiration levels also on the basis of the performance of the agents with which they interact. Since they interact with all firms via average profits, this is a model of social learning.

Aspiration States. Let’s first define for each pair of strategies the maximum and the minimum profits:

- $\pi_{min}(i, j) = \min[\pi_{i,j}, \pi_{j,i}]$
- $\pi_{max}(i, j) = \max[\pi_{i,j}, \pi_{j,i}]$

These two values, together with the average profits, tell us in which of four possible “aspiration states” a certain pair of strategies $\ell(i, j)$ is. We can, in fact, classify every pair of strategies on the basis of the firm-types that are going to experiment.

A pair of strategies is in aspiration state 1 if the profits attached to the two strategies of the pair are both greater than or equal to the average profits.

³As in Dixon (1996) and Palomino and Vega-Redondo (1996).

A pair of strategies is in aspiration state 2 (3) if the profits attached to strategy i (j) are lower than $\bar{\pi}_t$ and the profits attached to the other strategy j (i) are greater than or equal to $\bar{\pi}_t$.

Finally, a pair of strategies is in aspiration state 4 if the profits attached to both strategies are lower than $\bar{\pi}_t$.

We can now divide the set of all pairs of strategies \mathcal{L} in four partitions $\mathcal{L}_{z,t}$, $z \in \{1, \dots, 4\}$, according to the aspiration state of each pair at time t . More formally the elements of the four sets $\mathcal{L}_{1,t}, \dots, \mathcal{L}_{4,t}$ are defined in the following way:

$$\begin{aligned}\mathcal{L}_{1,t} &= \{\ell(i, j) \in \mathcal{L} : \bar{\pi}_t \leq \pi \min(i, j)\} \\ \mathcal{L}_{2,t} &= \{\ell(i, j) \in \mathcal{L} : \pi \min(i, j) = \pi(i, j) \leq \bar{\pi}_t < \pi(j, i) = \pi \max(i, j)\} \\ \mathcal{L}_{3,t} &= \{\ell(i, j) \in \mathcal{L} : \pi \min(i, j) = \pi(j, i) \leq \bar{\pi}_t < \pi(i, j) = \pi \max(i, j)\} \\ \mathcal{L}_{4,t} &= \{\ell(i, j) \in \mathcal{L} : \pi \max(i, j) < \bar{\pi}_t\}\end{aligned}$$

Clearly $\bigcup_z \mathcal{L}_{z,t} = \mathcal{L}$, and $\mathcal{L}_{s,t} \cap \mathcal{L}_{z,t} = \emptyset$ for every $s, z \in \{1, \dots, 4\}$ with $s \neq z$. Note that the partition is time dependent, since the set of pairs of strategies in a particular aspiration state at time t will depend on the average profits at t . Note that although the partition is defined in terms of *pairs*, it also implicitly partitions the set of duopolies, since each duopoly belongs to one and only one pair (i, j) . Hence we can define the four subsets $\mathcal{L}_{z,t}^d$:

$$\mathcal{L}_{z,t}^d = \{d \in [0, 1] : I(d, (i, j), t) = 1 \quad \text{and} \quad \ell(i, j) \in \mathcal{L}_{z,t}\}$$

Since there are L strategy pairs and four aspiration states, we define the the $4L \times L$ *aspiration state* matrix $\mathbf{\Lambda}_t$, which summarises the aspiration state z for each pair q at time t , being the following block diagonal matrix:

$$\mathbf{\Lambda}_t = \begin{bmatrix} \boldsymbol{\lambda}_{1,t} & 0 & \cdots & 0 \\ & \boldsymbol{\lambda}_{2,t} & & \vdots \\ 0 & & & 0 \\ \vdots & \ddots & & \ddots \\ 0 & \cdots & 0 & \boldsymbol{\lambda}_{L,t} \end{bmatrix} \quad (1)$$

The diagonal elements of $\mathbf{\Lambda}_t$ are the unitary vectors $\boldsymbol{\lambda}_{1,t}, \boldsymbol{\lambda}_{2,t}, \dots, \boldsymbol{\lambda}_{L,t}$, one for each strategy pair, whose only non-zero element is in the row corresponding to the aspiration state of the pair which the vector $\boldsymbol{\lambda}_{l,t}$ refers to. For instance, if at t the pair of strategies 1 is in the set $\mathcal{L}_{2,t}$, then, $\boldsymbol{\lambda}_{1,t} = [0 \ 1 \ 0 \ 0]'$. Hence, if the pair of strategies q is in learning state z , and v is the row index of the vector $\boldsymbol{\lambda}$ ($q = 1 \dots L$, $v = 1 \dots 4$), then:

$$\lambda_{v,q,t} = \begin{cases} 1 & \text{if } q \in \mathcal{L}_{z,t} \\ 0 & \text{otherwise} \end{cases}$$

Experimental states. Each duopoly can be classified in terms of the firms which are experimenting in it at time t . A firm is experimenting if it chooses to continue its existing strategy with a probability strictly less than 1. We can partition the population of duopolies $[0, 1]$ into four subsets \mathcal{E}_f , $f \in 1, \dots, 4$, which are mutually exclusive and exhaustive. These are:

$$\begin{aligned} \mathcal{E}_{1,t} &= \{d \in [0, 1] : \text{neither } i \text{ nor } j \text{ experiments at time } t\} \\ \mathcal{E}_{2,t} &= \{d \in [0, 1] : \text{only } i \text{ experiments at time } t\} \\ \mathcal{E}_{3,t} &= \{d \in [0, 1] : \text{only } j \text{ experiments at time } t\} \\ \mathcal{E}_{4,t} &= \{d \in [0, 1] : \text{both } i \text{ and } j \text{ experiment at time } t\} \end{aligned}$$

In absence of noise ($\delta = \beta = 0$) the partition of firms according to aspiration states and experimental states coincide: for duopoly d playing (i, j) , $d \in \mathcal{L}_f^d$ iff $\ell(i, j) \in \mathcal{E}_f$. However, if $\delta > 0$, any pair $q = \ell(i, j)$ can be in each of the four experimental states with a strictly positive probability whatever its aspiration state.

For any given pair $\ell(i, j) \in \mathcal{L}$, in experimental state z , there is a subset of pairs to which it can possibly move, the *destination set* $\mathcal{D}(z, (i, j)) \subseteq \mathcal{L}$:

$$\begin{aligned} \mathcal{D}(1, (i, j)) &= \{(i, j)\} \\ \mathcal{D}(2, (i, j)) &= \{(g, h) \in \mathcal{L} : \text{either } g = j \text{ or } h = j\} \\ \mathcal{D}(3, (i, j)) &= \{(g, h) \in \mathcal{L} : \text{either } g = i \text{ or } h = i\} \\ \mathcal{D}(4, (i, j)) &= \{\mathcal{L}\} \end{aligned}$$

The destination set reflects the fact that the destination pair must contain the current strategy of any firm not experimenting.

Corresponding to each aspiration state z , there is a 4×1 noise vector \mathbf{N}_z , giving the probability that a duopoly $d \in \mathcal{L}_{z,t}^d$ is in each of the 4 experimental states as a function of the noise parameters (δ, β) :

$$\mathbf{N}_z = \begin{bmatrix} \Pr(d \in \mathcal{E}_1 | d \in \mathcal{L}_{z,t}^d) \\ \Pr(q \in \mathcal{E}_2 | q \in \mathcal{L}_{z,t}^d) \\ \Pr(q \in \mathcal{E}_3 | q \in \mathcal{L}_{z,t}^d) \\ \Pr(q \in \mathcal{E}_4 | q \in \mathcal{L}_{z,t}^d) \end{bmatrix}$$

We have for strategy pairs $q \in \mathcal{L}_{1,t}$:

$$\mathbf{N}_1 = \begin{bmatrix} (1 - \delta)^2 & \delta(1 - \delta) & \delta(1 - \delta) & \delta^2 \end{bmatrix}'.$$

For strategy pairs $q \in \mathcal{L}_{2,t}$:

$$\mathbf{N}_2 = \begin{bmatrix} \beta(1-\delta) & (1-\beta)(1-\delta) & \beta\delta & (1-\beta)\delta \end{bmatrix}'.$$

For strategy pairs $q \in \mathcal{L}_{3,t}$:

$$\mathbf{N}_3 = \begin{bmatrix} \beta(1-\delta) & \beta\delta & (1-\beta)(1-\delta) & (1-\beta)\delta \end{bmatrix}'.$$

And finally, for strategy pairs $q \in \mathcal{L}_{4,t}$:

$$\mathbf{N}_4 = \begin{bmatrix} \beta^2 & \beta(1-\beta) & \beta(1-\beta) & (1-\beta)^2 \end{bmatrix}'.$$

These noise vectors represent the relationship in probabilistic terms between learning and experimental states. Together, the 4 noise vectors form the 4×4 matrix $\mathbf{N} = [\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}_3, \mathbf{N}_4]$:

$$\mathbf{N} = \begin{bmatrix} (1-\delta)^2 & \beta(1-\delta) & \beta(1-\delta) & \beta^2 \\ \delta(1-\delta) & (1-\beta)(1-\delta) & \beta\delta & \beta(1-\beta) \\ \delta(1-\delta) & \beta\delta & (1-\beta)(1-\delta) & \beta(1-\beta) \\ \delta^2 & (1-\beta)\delta & (1-\beta)\delta & (1-\beta)^2 \end{bmatrix} \quad (2)$$

The matrix \mathbf{N} is time invariant, because we have assumed that the values of δ and β are time invariant. In the case of no noise, then \mathbf{N} is the identity matrix ($\mathbf{N} = \mathbf{I}$).

We then have the $4L \times L$ *experimentation matrix* \mathbf{E}_t , which picks out the appropriate vector of experimentation probabilities for each pair q according to its aspiration state. \mathbf{E}_t is block diagonal with diagonal elements $\mathbf{N}\boldsymbol{\lambda}_{q,t}$:

$$\begin{aligned} \mathbf{E}_t &= \begin{bmatrix} \mathbf{N} & 0 & \cdots & 0 \\ & \mathbf{N} & & \vdots \\ 0 & & & 0 \\ \vdots & \ddots & & \\ 0 & \cdots & 0 & \mathbf{N} \end{bmatrix} \cdot \boldsymbol{\Lambda}_t \\ &= \begin{bmatrix} \mathbf{N}\boldsymbol{\lambda}_{1,t} & & 0 & \cdots & 0 \\ & \mathbf{N}\boldsymbol{\lambda}_{2,t} & & \ddots & \vdots \\ 0 & & & & 0 \\ \vdots & \ddots & & \ddots & \\ 0 & \cdots & 0 & & \mathbf{N}\boldsymbol{\lambda}_{L,t} \end{bmatrix} \quad (3) \end{aligned}$$

The Experimentation matrix \mathbf{E}_t is $4L \times L$ since each of the L pairs q can be in 4 experimental states f . In the no-noise case, since \mathbf{N} is the identity matrix, $\mathbf{E}_t = \boldsymbol{\Lambda}_t$.

Switching and Transition probabilities. We are now in a position to define the transition probabilities for strategy pairs. These are determined by the aspiration state of the pair; the corresponding experimental probabilities and the switching probabilities. We have already defined the first two, and will now introduce the switching probabilities.

Although it is individual firms that are changing their behaviour, it is easier to look at the probabilities of a pair of strategies switching to another pair as the effect of the change of one or both (or none) of the strategies that characterize that pair of strategies. Hence the switching probabilities of a given strategy pair will depend on the experimental state in which that pair is⁴.

We define the probability that a pair of strategies $\ell(i, j)$ switches to pair $\ell(g, h)$ at time t when it is in experimental state z as $\mu_{z,t}[(i, j), (g, h)]$, where of course $j \geq i$ and $h \geq g$. We write $(i, j) = (g, h)$ if and only if $i = g$ and $j = h$.

Formally, we define the 4×1 vector of switching probabilities at time t , $\mathbf{S}_t^{q,r}$ which contains the probabilities of switching from $q = \ell(i, j)$ to $r = \ell(g, h)$ in each of the four possible experimental states:

$$\mathbf{S}_t^{q,r} = \begin{bmatrix} \mu_{1,t}(q, r) \\ \mu_{2,t}(q, r) \\ \mu_{3,t}(q, r) \\ \mu_{4,t}(q, r) \end{bmatrix} \quad (4)$$

where of course $\mu_{z,t}(q, r) > 0$ only if $q \in \mathcal{D}(z, r)$ and 0 otherwise, and for each z :

$$\sum_{r \in \mathcal{L}} \mu_{z,t}(q, r) = 1$$

The *switching* matrix \mathbf{S}_t is the following $L \times 4L$ block matrix constructed out of the 4×1 vectors $\mathbf{S}_t^{q,r}$:

$$\mathbf{S}_t = \begin{bmatrix} \mathbf{S}_t^{1,1} & \mathbf{S}_t^{2,1} & \dots & \mathbf{S}_t^{L,1} \\ \mathbf{S}_t^{1,2} & \mathbf{S}_t^{2,2} & \dots & \mathbf{S}_t^{L,2} \\ \vdots & & \ddots & \vdots \\ \mathbf{S}_t^{1,L} & \mathbf{S}_t^{2,L} & \dots & \mathbf{S}_t^{L,L} \end{bmatrix} \quad (5)$$

The superscripts indicate the starting and the destination pair of strategies respectively.

⁴A more precise terminology is *conditional* switching probability, since the switching decision is brought about only if the firm has decided to experiment.

The $L \times L$ matrix of transition probabilities \mathbf{T}_t gives the transition probabilities $T_t^{q,r}$ of pair q to pair r at time t . This can be computed according to the following formula:

$$\mathbf{T}_t = \mathbf{S}_t \cdot \mathbf{E}_t \quad (6)$$

If we break down equation (6), we can see that \mathbf{E}_t picks out the appropriate vector of experimentation probabilities for each strategy pair at t , and is the $4L \times L$ block diagonal matrix with diagonal elements $\mathbf{N}\boldsymbol{\lambda}_{q,t}$. The block of resultant experimentation probabilities is then multiplied by each of the 1×4 vector $\mathbf{S}_t^{q,r}$ of switching probabilities for pair q to each pair r . Each element of \mathbf{T}_t is thus:

$$T_t^{q,r} = \mathbf{S}_t^{q,r} \cdot \mathbf{N} \cdot \boldsymbol{\lambda}_{q,t} \quad (7)$$

where $T_t^{q,r}$ is a scalar, the resultant transition probability between two strategy pairs at time t .

3 Specific Switching rules

We can now introduce the three switching technologies adopted by the firms when they experiment which we explore in this paper: random-switching, best-response and imitation.

Random switching. According to this technology the firms that are experimenting simply randomly switch to another strategy. Since all the strategies are equally likely to be selected, the probability for a firm that is experimenting of selecting any other strategy (included the one that it was adopting) is $\frac{1}{K}$.

Consider duopoly d with firms choosing strategies (i, j) . If $d \in \mathcal{E}_1$ we have (dropping the time subscripts for convenience):

$$\mu_1[(i, j), (g, h)] = \begin{cases} 1 & \text{if } (g, h) = (i, j) \\ 0 & \text{if } (g, h) \neq (i, j). \end{cases}$$

For $d \in \mathcal{E}_2$:

$$\mu_2[(i, j), (g, h)] = \begin{cases} 1/K & \text{if either } g = j \text{ or } h = j \\ 0 & \text{otherwise.} \end{cases}$$

For $d \in \mathcal{E}_3$:

$$\mu_3[(i, j), (g, h)] = \begin{cases} 1/K & \text{if either } g = i \text{ or } h = i \\ 0 & \text{otherwise.} \end{cases}$$

And finally, for $d \in \mathcal{E}_4$:

$$\mu_4[(i, j), (g, h)] = \begin{cases} 1/K^2 & \text{if } g = h \\ 2/K^2 & \text{if } g \neq h. \end{cases}$$

Note that there is an important difference between states 2, 3 and state 4: if a particular pair of strategies is in the set \mathcal{E}_4 , there is a positive probability that it will switch to any and all pairs of strategies in \mathcal{L} (since $\mathcal{D}(4, q) = \mathcal{L}$ for all q). There is no pair of strategies $q \in \mathcal{L}$ that has a zero probability of being reached. In states 2 or 3, where only one firm experiments, there is a restriction on the pairs of strategies it can reach (this restriction comes from the fact that the non-experimenting firm does not change).

Under RS, the switching matrix does not vary over time, $\mathbf{S}_t = \mathbf{S}$ for all t . Hence the only element to vary in the transition matrix is \mathbf{A}_t , which reflects the changing aspiration states of the pairs as average profits vary.

Imitation. We now consider a different learning rule based on imitation. In societies the behaviour of an individual is influenced and influences the behaviour of other members of the society. Therefore within societies an individual may learn mainly through observation and imitation. Imitation, intended as the act of mimicking the behaviour of another individual has as one of its main advantages the fact that it does not require any particular ability or skill and any prior knowledge of others' payoff function (see Schlag (1996b,a); Weibull (1995)).

The probability at time t of imitating the strategy of another firm depends on the proportion of firms adopting that strategy at time t . The higher the proportion of a certain strategy, the greater the probability of that strategy of being imitated. Hence switching behaviour will evolve with the population proportions $\hat{P}_t(i)$ where $i = 1 \dots K$. In this case the switching probabilities are the following:

for duopolies $d \in \mathcal{E}_1$:

$$\mu_{1,t}[(i, j), (g, h)] = \begin{cases} 1 & \text{if } (g, h) = (i, j) \\ 0 & \text{if } (g, h) \neq (i, j). \end{cases}$$

For duopolies $d \in \mathcal{E}_2$:

$$\mu_{2,t}[(i, j), (g, h)] = \begin{cases} \hat{P}_t(g) & \text{if } h = j \\ \hat{P}_t(h) & \text{if } g = j \\ 0 & \text{otherwise.} \end{cases}$$

For duopolies $d \in \mathcal{E}_3$:

$$\mu_{3,t}[(i,j), (g,h)] = \begin{cases} \hat{P}_t(g) & \text{if } h = i \\ \hat{P}_t(h) & \text{if } g = i \\ 0 & \text{otherwise.} \end{cases}$$

For duopolies $d \in \mathcal{E}_4$:

$$\mu_{4,t}[(i,j), (g,h)] = \begin{cases} \hat{P}_t(g) \hat{P}_t(h) & \text{if } g = h \\ 2\hat{P}_t(g) \hat{P}_t(h) & \text{if } g \neq h. \end{cases}$$

Best-response. With the best response (BR) dynamic the firm that decides to experiment chooses the strategy that maximizes its payoff against its competitor's current strategy. Since we are dealing with a finite strategy set, it is simple to list the best response(s) to each strategy. Let us define for each strategy i the set of best responses $B(i)$:

$$B(i) = \arg \max_{j=1,\dots,K} \pi_{ji}$$

In effect we are choosing the largest element(s) along the i^{th} column of the payoff matrix Π . Let us denote the cardinality of $B(i)$ as $\#B(i)$. Clearly, best responses are time independent. The switching probabilities are the following (where we drop the time subscripts):

for duopolies $d \in \mathcal{E}_1$:

$$\mu_1[(i,j), (g,h)] = \begin{cases} 1 & \text{if } (g,h) = (i,j) \\ 0 & \text{if } (g,h) \neq (i,j). \end{cases}$$

For duopolies $d \in \mathcal{E}_2$:

$$\mu_2[(i,j), (g,h)] = \begin{cases} \frac{1}{\#B(j)} & \text{if either } g = j \text{ or } h = j \\ 0 & \text{otherwise.} \end{cases}$$

For duopolies $d \in \mathcal{E}_3$:

$$\mu_3[(i,j), (g,h)] = \begin{cases} \frac{1}{\#B(i)} & \text{if either } g = i \text{ or } h = i \\ 0 & \text{otherwise.} \end{cases}$$

And finally, for duopolies $d \in \mathcal{E}_4$:

$$\mu_4[(i,j), (g,h)] = \begin{cases} \frac{2}{\#B(i)\#B(j)} & \text{if } g, h \in B(i) \cap B(j) \\ \frac{1}{\#B(i)\#B(j)} & \text{if } (g,h) \in B(i) (\Delta B(j)). \end{cases}$$

The BR switching rule is rather different from the RS and IM rules. As in the case of RS, the switching probabilities for each pair in each aspiration

state are constant over time ($\mathbf{S}_t = \mathbf{S}$ for all t). However, the switching probabilities will be zero for all strategy pairs which are not best responses for the experimenting firms. Unlike RS and IM, the BR rule depends on the strategic nature of the duopoly in which the experimenting firm finds itself. However, it leads to some strange learning behaviour. For example, consider the case of no noise. If the current strategy chosen by the firm earns below average profits, then if it is a best response the firm will “experiment” by choosing the same strategy again and again (so long as the other firm does not switch). The BR rule restricts experimentation, and can lead to a firm becoming “locked in” to the current strategy even when it is below aspiration.

3.1 Population dynamics

The state of the economy at time t is summarised by the $L \times 1$ vector $\mathbf{P}_t = [P_t(1), \dots, P_t(L)]'$, where $P_t(q)$ gives the proportion of duopolies with strategy pair $q \in \mathcal{L}$ at time t . We can write the population dynamics as:

$$\mathbf{P}_{t+1} = \mathbf{T}_t \mathbf{P}_t \quad (8)$$

This equation is a non-stationary Markov process, since in general the transition matrix \mathbf{T} varies with t . The transition matrix will vary over time as pairs change aspiration state ($\mathbf{\Lambda}_t$ varies), and in addition the switching matrix \mathbf{S}_t may vary. The causal mechanism is that the current \mathbf{P}_t determines the average profit rate $\bar{\pi}_t$, which determines the aspiration state matrix $\mathbf{\Lambda}_t$, which then determines the proportions of each location q experimenting, and the switching matrix the proportions of experimenting pairs to $r \in \mathcal{L}$. In general, such processes are hard to analyse. However, in the absence of noise ($\delta = \beta = 0$) the dynamics can be solved analytically under certain assumptions. In order to clarify this, let us make the following definitions:

$$\begin{aligned} \text{Sym} &\subset \mathcal{L} \text{ where } (i, j) \in \text{Sym} \text{ iff } \pi_{i,j} = \pi_{j,i} \\ S &= \arg \max_{\text{Sym}} \pi_{i,j} \\ \Pi S &= \arg \max_{\text{Sym}} [(\pi_{i,j} + \pi_{j,i})/2] \\ \text{maxav} &= \arg \max_{\mathcal{L}} [(\pi_{i,j} + \pi_{j,i})/2] \end{aligned}$$

Sym is the set of payoff symmetric pairs; S is the set of elements of Sym which maximize joint payoff; ΠS is the set of pairs that maximizes the joint payoff over pairs in Sym; maxav maximizes the joint payoff over all pairs \mathcal{L} . We can make two assumptions, as in Dixon (1996):

A1 : $\Pi S = \max_{\mathbf{a}} \mathbf{a} \cdot \mathbf{S}$

A2 : There exists $\varphi > 0$ such that for all $t = 0 \dots \infty$, and all $q \in \mathcal{L}$, $\mathbf{S}_t^{q,r}$ in (4) satisfies:

$\mu_{z,t}(q,r) > \varphi$ for $z = 2 \dots 4$ and $r \in \mathcal{D}(z,q)$ if $z = 1$, $\mu_{1t}(q,r) = 1$ iff $q = r$ and $\mu_{1t}(q,r) = 0$ otherwise.

Theorem 1 *A1 and A2, $\beta = \delta = 0$. If $P_0(S) > 0$, then as $t \rightarrow \infty$, $P_t(S) \rightarrow 1$.*

The Cournot model we use for the simulations satisfies A1. The switching rules we have employed may or may not satisfy A2, which states that all of the feasible destination elements of the switching matrix \mathbf{S}_t must be greater than φ in learning states 2–4. This is generally violated by BR dynamics since non-best responses have zero probabilities. With IM, since $P_t(q)$ may go to zero, A2 can also be violated. Hence we need to simulate the system in equation (8).

4 Results: Simulating the The Cournot Duopoly Model

In this section we present our simulation results with the three different switching rules RS, IM and BR. Since RS can be interpreted as pure noise, we are also considering mixtures of random switching with the other two rules to represent “noisy” versions of those rules. These different rules express different ways of computing the matrix \mathbf{S}_t of switching probabilities. We will also consider for each process different levels of noise in the experimentation decision, δ and β . These imply particular specifications of the noise matrix \mathbf{N} used to calculate the of experimentation probabilities. In section 4.1 we consider the noiseless ($\delta = \beta = 0$) version of all three switching rules. Then in sections 4.2-4 we consider the different switching regimes.

We have adopted the simplest possible Cournot model as an example of the general model presented in the previous section. The market price p is a linear function of the output produced by the two firms at the location:

$$p = \max[0, 1 - x_i - x_j]$$

where $x_i, x_j \geq 0$ are the outputs of the two firms. We assume that there are no production costs. The price can be interpreted as the price net of constant average production costs so that the model can allow for constant

Table 1: Cournot Duopoly: Reference Points

outcome	output	profit per firm
Cournot-Nash Equilibrium	$x = y = \frac{1}{3}$	0.1111
Joint-Profit Maximum	$x = y = \frac{1}{4}$	0.1250
Stackelberg	$x = \frac{1}{2}, y = \frac{1}{4}$	0.1250, 0.0625
Walrasian	$1 - x - y = 0$	0

unit production costs. In this case, the profits of the x_i -firm are: $\pi_i = x_i \cdot p$. We identify some of the key reference points of the model in table 1.

We have reported the profit levels to 4 s.f. In this sort of Cournot-Nash model, the outputs of the two firms are strategic substitutes (the best-response functions are downward sloping), and the payoff of each firm is decreasing (when positive) in the output of the other firm. In order to generate a finite strategy set from this model, we simply construct a grid of outputs, so that each strategy is an output level. The payoff matrix Π gives the payoffs $\pi_{i,j}$ to firm producing output levels $i = 1 \dots K$ when the other firm is producing output levels $j = 1 \dots K$.

The simulations we have made are based on two specifications of the set of permissible outputs. In both cases the set of outputs (firm types or strategies) is generated by a grid search over the interval $[0.1, 0.6]$. In specification 1 there are 11 firm types and hence 66 pairs, and the granularity of the grid is 0.05; in specification 2, the granularity is 0.025, resulting in 21 firm types, and 231 pairs. In both cases we perturb the grid slightly to ensure that the Cournot firm (output 0.3333) is included. In section 4.1 we use the large set (specification 2), and for the rest of the results we use specification 1 unless otherwise specified. We report all outputs and profits to 4 s.f. unless otherwise stated.

4.1 The case of no noise in the decision to experiment ($\delta = \beta = 0$)

From Dixon (1996), we know that under the assumption that the switching probabilities are bounded away from zero, the learning process outlined in this paper will lead asymptotically to all duopolies being JPM, with a level of profits 0.125. However, the Theorem tells us little or nothing about the path to this position. It is useful, therefore to consider how the population evolves when there is no experimental noise. In fact, of the three switching regimes that we consider, only the random switching rule strictly satisfies

the assumptions of the theorem, since imitation and best response can or do involve some switching probabilities being zero (or arbitrarily close to it). However, as we shall see this appears not to matter in the case of imitation. In all three cases, the initial position was the uniform distribution, so that all pairs of outputs are represented equally.

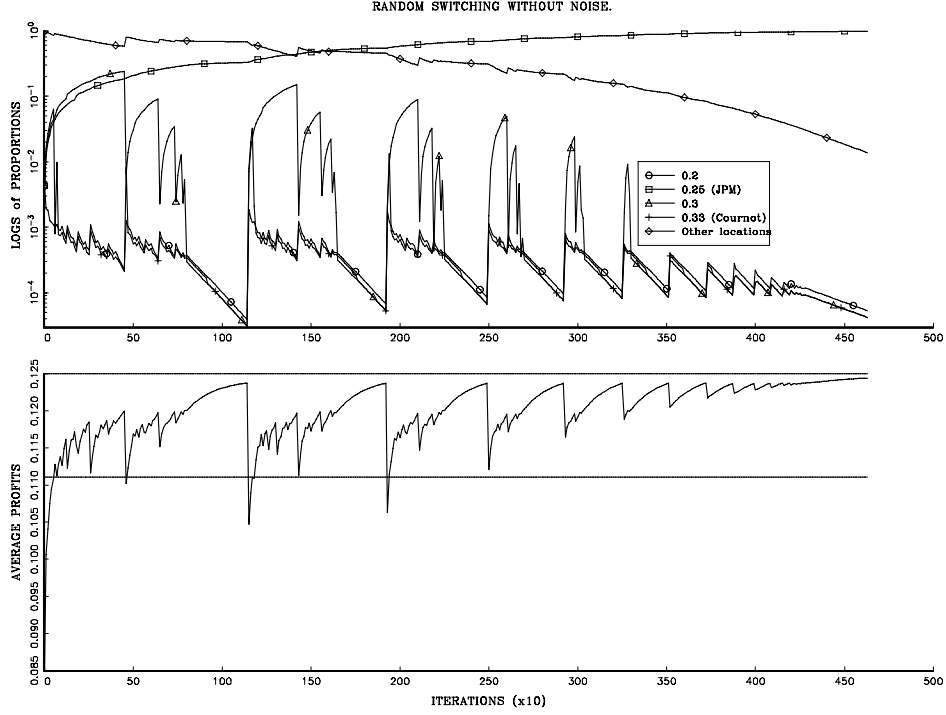


Figure 1: Evolution under RS.

Firstly, let us compare random switching and imitation. In the lower half of fig. 1 we have the evolution of average profits under RS, and in the lower half of fig. 2 the average profits under IM. Above them we have the population shares of four of the 231 pairs. These are all symmetric pairs with outputs as in table 2.

Note that the scale in Figures 1 and 2 is logarithmic. We make the following observations.

- 1 Both switching processes converge on the collusive outcome⁵, with the proportion of locations with JPM firms tending to unity. Furthermore, since the JPM location is an absorbing state, the share of JPM locations

⁵Note, however, that for IM to converge to JPM, it is necessary that the initial proportion of firms choosing 0.125 needs to be strictly positive, otherwise the probability of switching to 0.125 will always be 0.

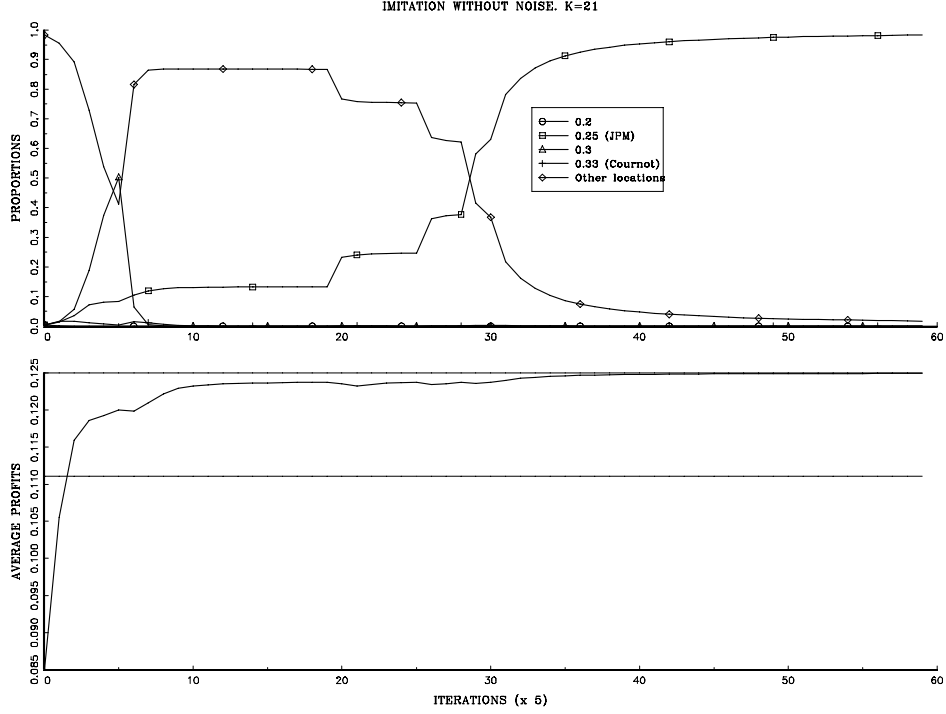


Figure 2: Evolution under IM.

is monotonic. However, under IM the convergence is much quicker: within 200 periods, the share of JPM locations is 95.26%, whereas with RS, at the same stage, the proportion is only 11.10%. This is a very stark difference which arises because the probability of firms that experiment arriving at the JPM pair evolves very differently. Under IM, the probability that a firm which is experimenting chooses 0.25 is equal to the share of this output in the population. As the JPM location becomes more common, this probability increase, thus if 90% of firms are 0.25, then a pair in aspiration state 4 has a 0.81 transition probability of switching to the JPM location. However, under RS, the probability of a pair in aspiration state 4 switching to JPM is always the same: $1/K^2$ (0.002267 when $K = 21$ as in our first specification of the set of permissible outputs).

- 2 If we compare the path of average profits, we can see that they are very different. Under RS, average profits are subject to discrete jumps and the time path is highly non-monotonic. The reason for this is that there are discrete and discontinuous changes in the transition matrix as the level of profits changes. Starting from an initial value, all locations

Table 2: four reference pairs

type	output
JPM	0.20
	0.25
	0.30
Cournot	0.3333

with both firms earning above average will be temporary “absorbers”, in the sense that there will be net inflows from the locations where one or both firms are earning below average. However, at certain times the level of average profits equals the profits of the least profitable firm at locations with a particular pair of outputs. As soon as the average profit rises, all of these locations switch from aspiration state 1 to state 2, 3 or 4. This will cause a discrete fall in profits, since under RS all outputs are equally likely, and even unprofitable pairs of outputs will result.

- 3 Under IM, the path of profits is smooth and only slightly non-monotonic. This behaviour results from the fact that in learning states 2, 3 or 4, the pair is more likely to transit to an above average pair.
- 4 In both cases, in the early stages, JPM is not the most common pair. In fact, both the symmetric locations 0.3 and 0.3333 (Cournot) initially emerge as more common than JPM. This occurs because in the Cournot model, the firm with the larger market share tends to earn more profits. Thus both 0.3 and 0.3333 have inflows from asymmetric pairs. For example, if we have the pair (0.25, 0.3333), the profits of the two firms are (0.1042, 0.1388). Clearly, the 0.25 firm will experiment at low levels of profit, and some of these firms will choose either 0.3 or 0.3333. Thus the symmetric locations 0.3 and Cournot will absorb firms from these asymmetric locations. The argument does not hold for symmetric locations with larger outputs, since at these the levels of profits are too low (for example at the symmetric location 0.4 the payoff is 0.08, as compared to 0.1111 at Cournot and 0.12 at 0.3) and they are not “absorbing” for long enough.

4.2 Noise in the experimentation decision with random switching.

The introduction of noise into the decision to experiment substantially alters the dynamics of the system. Most importantly, if $\delta > 0$ the JPM location is no longer absorbing: the share of locations at JPM can decline. We ran simulations which started from a uniform distribution of locations over pairs. We considered different magnitudes of noise, and ran simulations where $\delta = \beta > 0$, and simulations with just $\delta > 0$ ($\beta = 0$). The reference magnitudes for noise were 0.1, 0.01, 0.001 and 0.0001 (we found that noise higher than 0.1 made little difference). With noise, the simulations take longer to run, and we used the small grid with 11 firm types (66 pairs). In Figures 3 and 4 we depict the shares and average payoff series for low noise and high noise with δ and β noise.

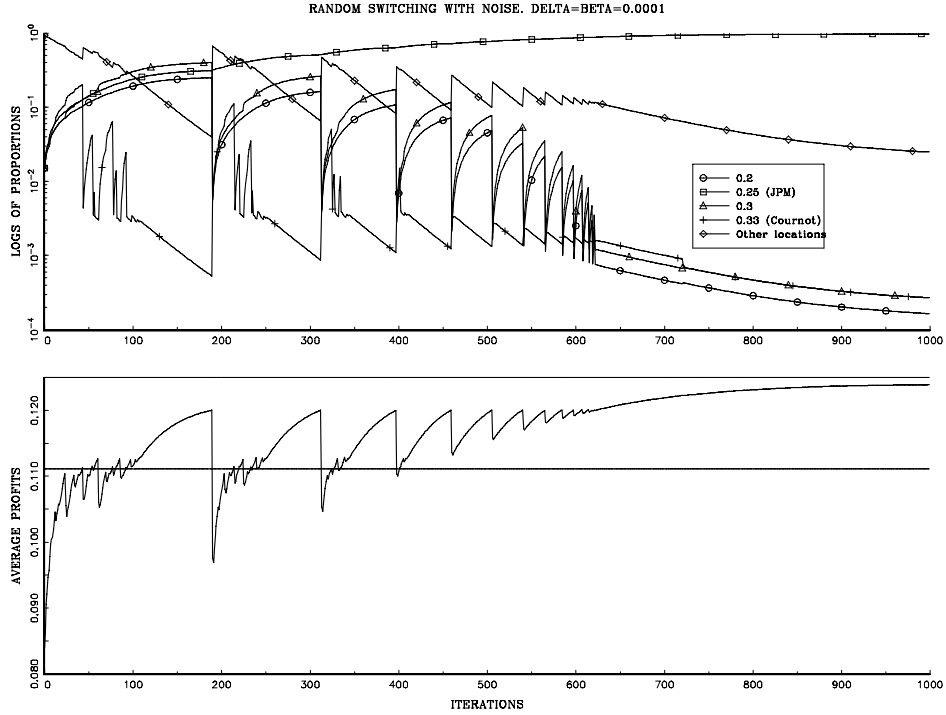


Figure 3: Evolution under RS ($\delta = \beta = 0.0001$).

With a small level of noise, the time series are very similar to the no noise case. In particular, the proportion of JPM locations remains monotonic, and the average profit is highly non-monotonic. This means

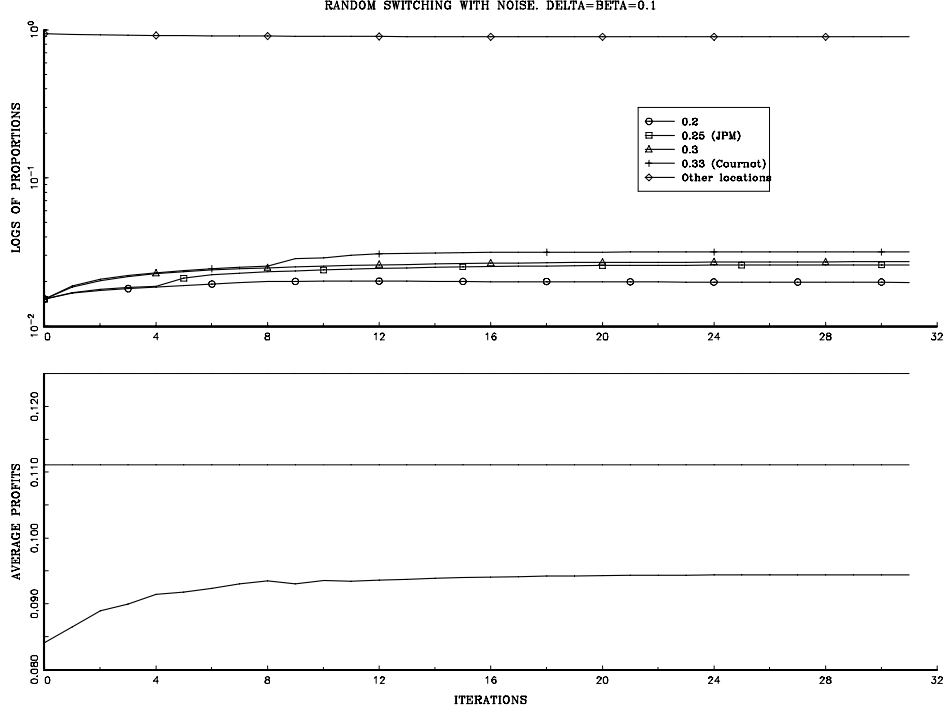


Figure 4: Evolution under RS ($\delta = \beta = 0.1$).

that the no-noise case is robust. However, with high levels of noise, the picture is rather different. The main differences are:

- 5 The series for proportions and profits converged after a few periods (34), and they are much smoother.
- 6 The limiting level of profits is low: 0.0944, reflecting that the average market is more competitive than the Cournot outcome.
- 7 A lot of pairs survive. The 4 reference pairs together have a share of only 10.13%. However, the Cournot pair has the largest population share (3.1%): this is followed closely by the JPM location and the other reference locations.

How does the magnitude of the noise level influence the outcome of the population dynamics? In Fig. 5, we show the end distribution at the different magnitudes of noise. We note that:

- 8 The share of JPM (0.25) declines with noise, as do all of the reference locations.

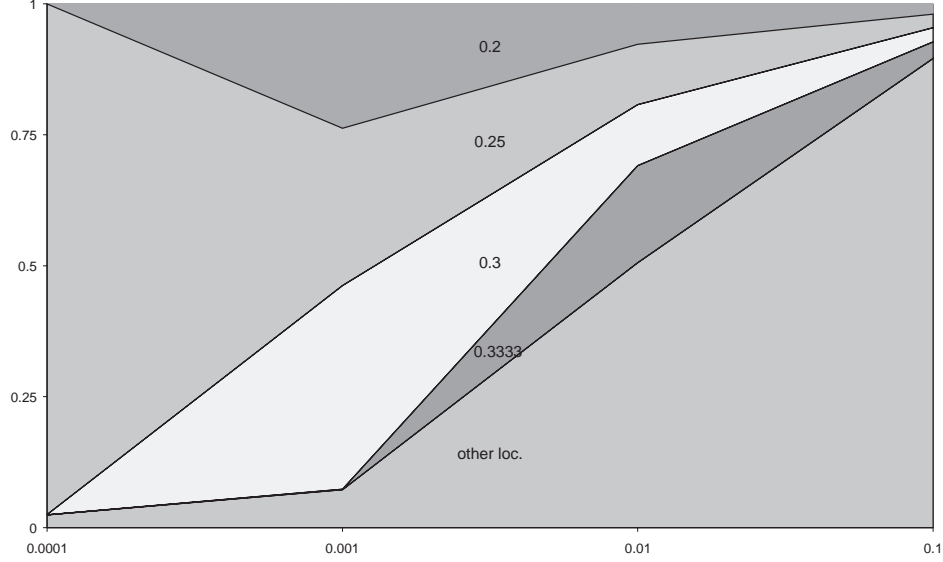


Figure 5: End distributions of RS noisy evolutionary processes.

- 9 JPM is not the largest proportion: Cournot (0.3333) and 0.3 both have larger shares for magnitudes above 0.0001.
- 10 With noise equal to 0.01 we discovered a limit cycle of 4 periods. This is depicted in Fig. 6. We can see that the Cournot and JPM pairs cycle slightly.

4.3 Imitation with noise

The presence of noise with IM introduces a very different population dynamic to the model. This is because the switching probabilities now depend on the population proportions. Most importantly, this means that the initial position can be important in determining the final outcome.

Let us first consider the simulations for different levels of noise with the same initial position of a uniform distribution. We show the case of $\delta = \beta = 0.0001$ in Fig. 7, and $\delta = \beta = 0.01$ in Fig. 8. In these (and the intermediate case), the pattern is the same.

- 11 The proportion of JPM locations is non-decreasing. However, in the initial stages, two other pairs have the largest population share: the Cournot and the 0.3 pairs. The limiting share of JPM is close to 1

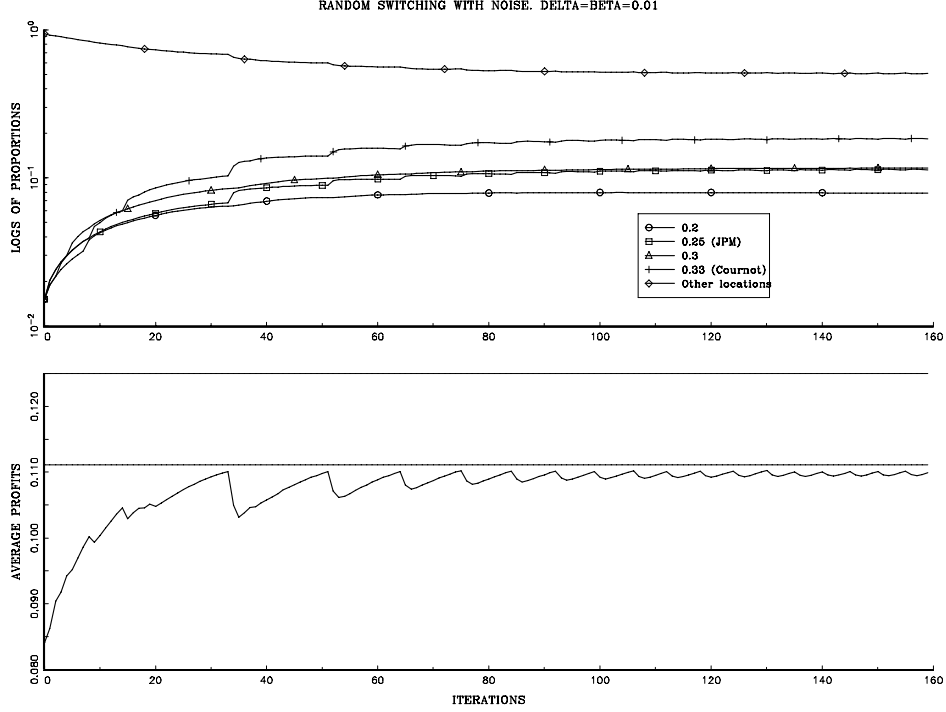


Figure 6: Random switching with noise ($\delta = \beta = 0.01$).

- 12 Average profits are not monotonic when $\delta = \beta = 0.01$, but almost so. They are monotonic when $\delta = \beta = 0.0001$. In both cases, the average profits tend to a figure close to 0.125.

Both of these observations are consistent with the no-noise case in section 4.1: the Cournot and 0.3 pairs decline rapidly after the level of the average payoff exceeds their own. However, with high noise 0.1, the system looks rather different. In Fig. 9 we give the time series of payoff and proportions.

- 13 With high noise, the 0.3 location ends up with a share close to unity: this is at a level of competitiveness in between the JPM and Cournot pairs, with an average profit of 0.12.

In order to better understand the behaviour of the system under IM with high levels of noise, we restricted our attention to a limited range of outputs: 0.25, 0.3, and 0.3333, with 6 possible pairs. These are the only output levels that seem to survive with imitation. Using this restricted set of outputs, we were able to explore the issue of the initial position systematically. We depict the state of the economy in terms

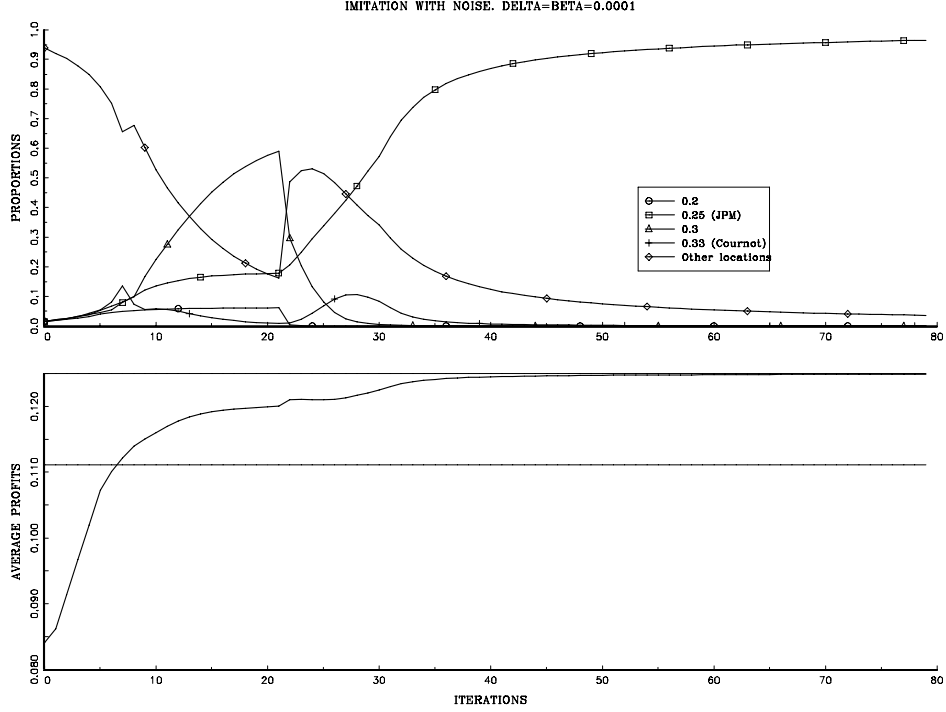


Figure 7: Imitation with noise ($\delta = \beta = 0.0001$).

Table 3: Proportion of pairs

pair	1,1	1,2	1,3	2,2	2,3	3,3
proportion	\hat{P}_1^2	$2\hat{P}_1\hat{P}_2$	$2\hat{P}_1\hat{P}_3$	\hat{P}_2^2	$2\hat{P}_2\hat{P}_3$	\hat{P}_3^2

of the shares of the 3 output levels: the simulations are run with the 6 pairs, and what we see on the unit triangle is the result. With our imitation rule, the switching probabilities are related to the proportions of each type, which is summarised by the position in the unit simplex. We ran 1275 simulations for each level of noise. This was generated by a grid on the unit triangle: the granularity of the grid was 0.02 ($\frac{1}{50}$). The initial position in terms of firm proportions is then mapped onto pairs by the formula depicted in Table 3.

This distribution of firm types over pairs is thus proportional to the proportions of firm types. The results are depicted in Fig. 10a-d⁶. In Fig. 10a-d there are four different levels of δ noise ($\beta = 0$).

⁶The convergence criterion used was $\mathbf{P}_t - \mathbf{P}_{t-1} \leq 0.0000316$.

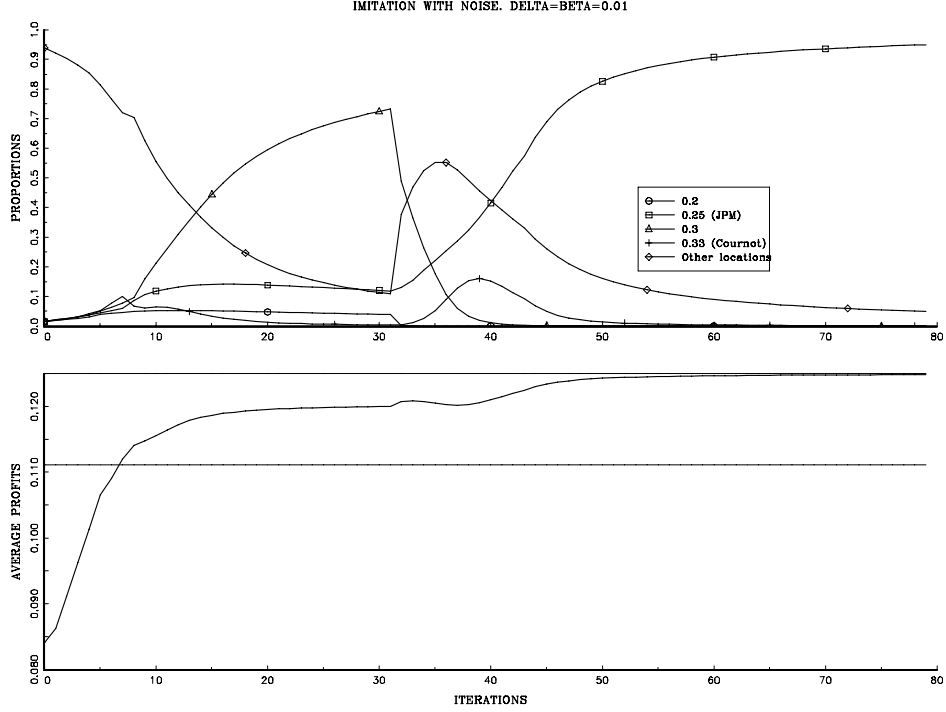


Figure 8: Imitation with noise ($\delta = \beta = 0.01$).

- 14 Without noise, we know that there are two attractors⁷. Thus, if we start from distributions along the bottom of the simplex, the attractor will be the the 0.3 vertex (the right one). However, if the initial proportion of 0.25 firms is strictly positive, then the attractor is all firms being 0.25 (the top 0.25 vertex).
- 15 With noise, we found that there were two types of attractors. In the first case, most firms where choosing 0.25; in the second case most firms were choosing 0.3. Due to noise, it is not possible to have all firms choosing the same strategy. The levels of δ noise are 0.0001, 0.001, 0.01 and 0.1. The dark area at the bottom is the basin of attraction for the mainly 0.3 attractor. As we can see, it gets bigger as δ increases in size. This explains the results with the full set of firms (ix): with $\delta = 0.1$ the uniform initial distribution is within the basin of attraction of the mainly 0.3 end distribution.

⁷With 3 strategies there are 6 pairs as depicted in table 3. If the initial proportion of JPM locations is strictly positive then JPM is the attractor. If the initial proportion of JPM locations is zero, and the proportion of pair 0.3 is strictly positive, then the attractor is the pair 0.3

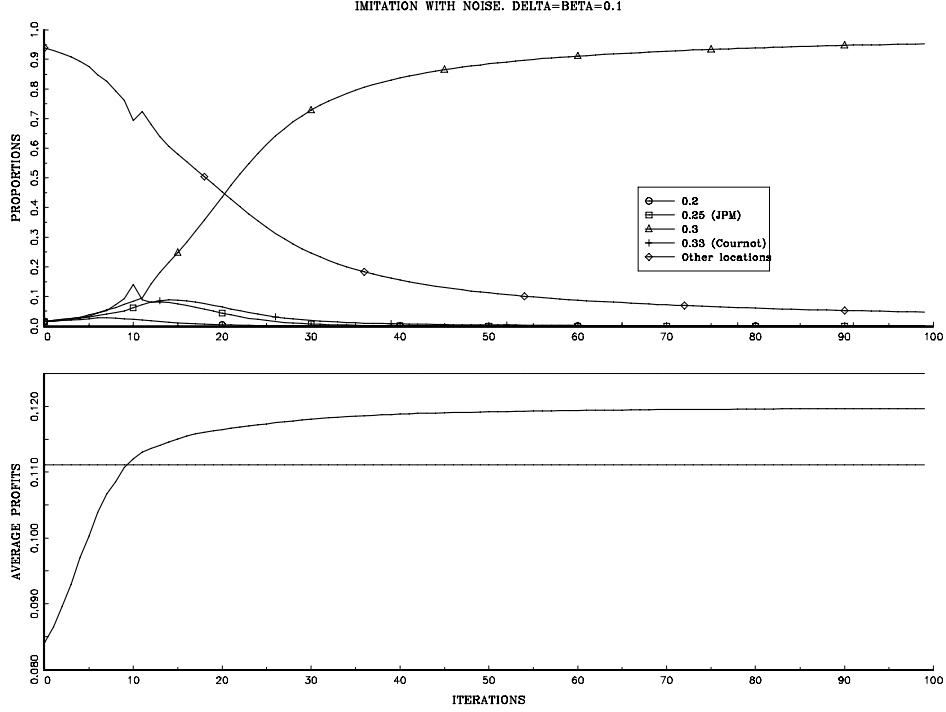


Figure 9: Evolution under IM with noise ($\delta = 0.1$).

The dependence of the limiting distribution on the initial position with IM is due to the fact that the switching probabilities depend on the point you are at in the unit simplex. This sensitivity to the initial position is lacking in the case of BR and RS. Once a level of output becomes predominant, it is likely that it will increase, since a large proportion of switching firms will switch towards it.

4.4 Best Response (BR) dynamics

In table 4, we list the strategies used in the simulation and the best response(s) to them for the case where there are 11 firm types.

Note that in three cases the best response is not unique, there being two. In this case, we assume that the switching probabilities are equal (0.5).

We have run one simulation without noise with the large set of firm types (21). We have also run 8 simulations with the 4 magnitudes of noise, one set with δ noise, and one with both types of noise.

In the case of absence of noise, the results were in contrast to the case of IM and RS. We found that the simulations converged after only 7 iterations, and that the distribution of firm types and locations pairs was quite different

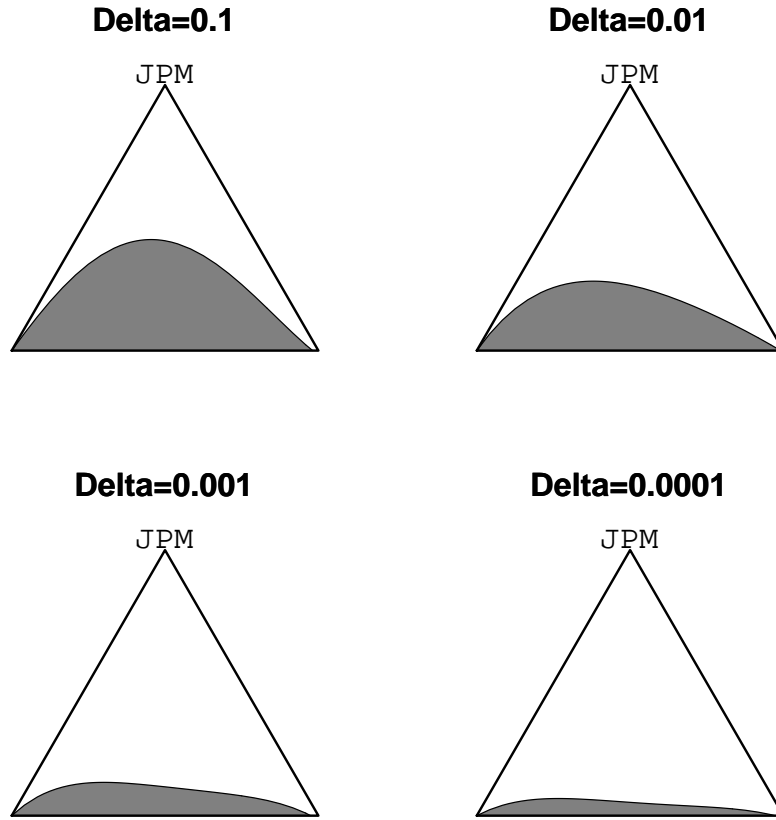


Figure 10: Basins of attraction of the IM rule.

to the other two cases. Perhaps the main reason for this difference is that the noiseless BR dynamic places a zero probability on switching to all locations except the current best responses. In Fig. 11, we depict the evolution of proportions and average payoff under noiseless BR.

- 16 Without noise, the most common pair is the Cournot pair, with a limiting share of 0.2338 and the limiting level of average profits is 0.1088. Those locations where both firms earn above 0.1088 never experiment, and their population shares remain constant.
- 17 Unlike RS and IM, it is possible for asymmetric pairs to survive in the limiting distribution. The reason is that one firm can be earning above average profits, and the low profit firm is choosing the best response to the high output/high profit firm. Consider the example of the pair (0.3, 0.4). This pair survives with a proportion 4.33%. The high output firm earns 0.12, and so does not experiment: the low output firm earns

Table 4: Best Response correspondence.

Strat.	Quantity	Best Response	Strat.	Quantity	Best Response
1	0.1	8	7	0.4	5
2	0.15	7,8	8	0.45	4,5
3	0.2	7	9	0.5	4
4	0.25	7	10	0.55	3,4
5	0.3	6	11	0.6	3
6	0.3333	6			

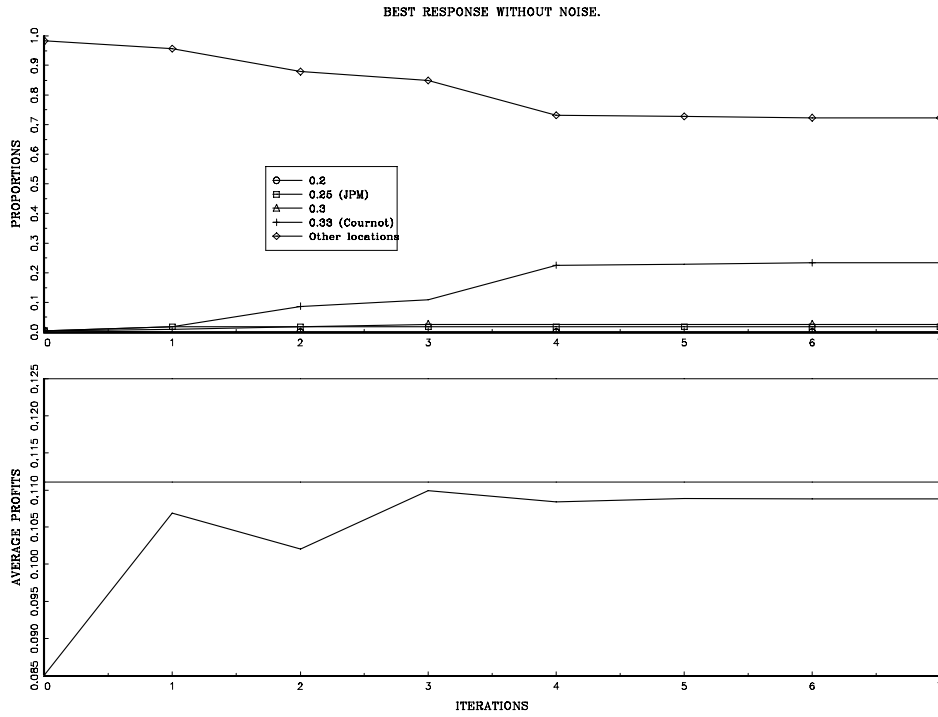


Figure 11: Evolution under BR.

only 0.09 which is below average, but since the low output is the best response to the high output, it does not change its output. In the limiting distribution, the share of asymmetric output pairs is 65.37%. BR is a very limited form of experimentation. Its restrictive nature leads to this odd form of “locking” to low profit, but best response strategies.

With experimental noise, however, the system behaves very differently to

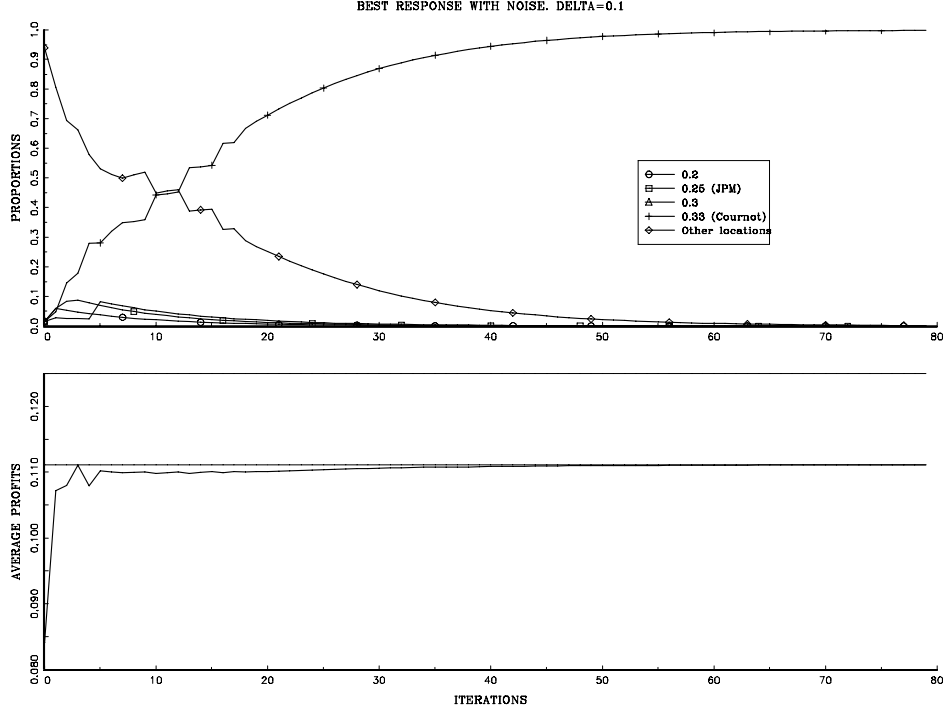


Figure 12: Evolution under BR with noise ($\delta = 0.1$).

the no noise case. So long as $\delta > 0$, there is a strictly positive probability that the firms with above average profits will switch to their best responses to the other firm. We depict the simulations for $\delta = 0.1$ and $\delta = 0.001$ in Figures 12 and 13. In all cases we have looked at, we found:

- 18 So long as $\delta > 0$, the Cournot pair is the one which has the lion's share of the population in the limiting distribution (the value of β is unimportant). This is certainly very close to unity with large values of δ , and possibly also with smaller values⁸.

The reason behind [18] is fairly intuitive. If all firms (irrespective of current profits) are changing their strategy in accord with the BR dynamic, then the attractor is the Cournot pair. When $\delta > 0$ all firms are switching with a strictly positive probability, and the result is much the same (although the speed of convergence is of course slower when δ is smaller). If $\delta = \beta = 1$ all firms always choose their best response, which leads rapidly to the Cournot outcome, as in the standard Cournot stability analysis.

⁸With $\delta = 0.0001$ the simulation had not converged after 6000 iterations.

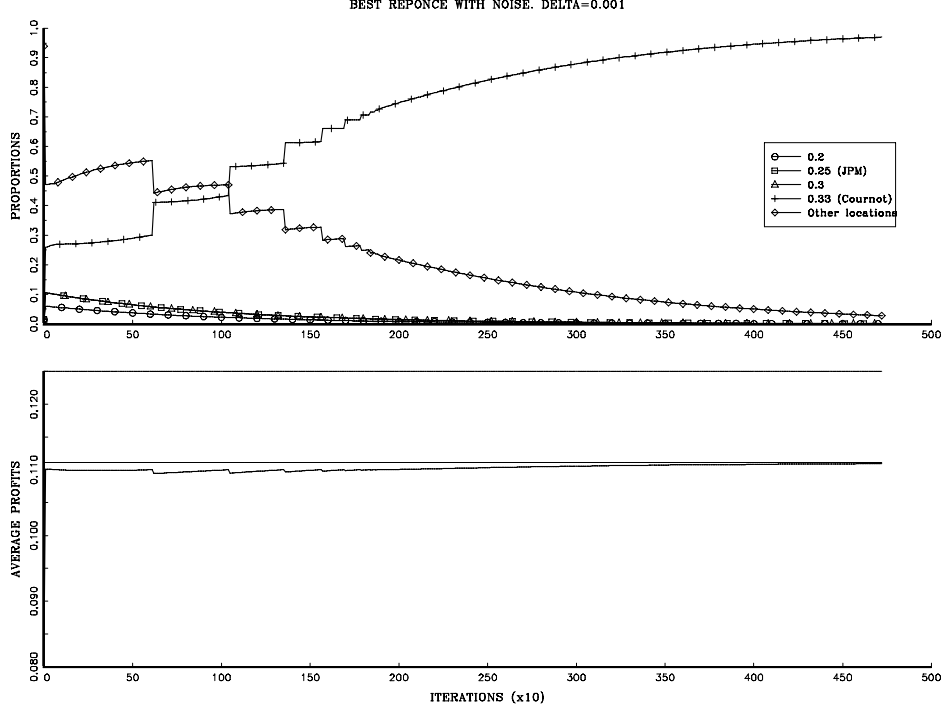


Figure 13: Evolution under BR with noise ($\delta = 0.001$).

4.5 Mixed switching rules.

In this section we consider mixtures of switching rules. In particular, we consider the two types that result from mixing IM and BR with RS. More specifically, we take a convex combination of the IM switching probability with the RS switching probability, with γ being the weight of the RS. This is an interesting exercise, since we can think of RS as representing “noise” in the switching process. Hence whereas δ and β represent noise in the experimentation decision (the \mathbf{N} matrix), γ represents the noise in the switching process represented by \mathbf{S}_t . We have run simulations: firstly, with no experimental noise ($\delta = \beta = 0$), we have 3 values for γ (0.25, 0.5, 0.75); secondly for $\gamma = 0.5$ we introduce different levels of experimental noise $\delta = \beta = 0.0001, 0.001, 0.01, 0.1$, and also with just the δ -noise. As in the cases of RS and IM, we found that β noise does not affect the shape of things.

In Figures 14 - 16 we depict the mixed rule with no noise, and different values of γ . As we can see:

- 19 No noise $\delta = \beta = 0$, IM and RS. For all three values of γ , we found that the pattern of the evolution of proportions was roughly the same. After an initial period where the Cournot location began to predominate,

the level of average profits passes 0.111 and from then on the JPM pair becomes the most common and tends to unity (as predicted by the Theorem). However, the evolution of profits and proportions is smoother and quicker when γ is larger (IM is more predominant).

In figures 17 - 19 we depict the end distribution of outputs for different values of δ noise ($\delta = 0.1, 0.01, 0.001$) with $\gamma = 0.5$.

- 20 Introducing experimental noise alters the evolution of the population, with low noise (fig. 19), the pattern is similar to the no-noise case, with JPM emerging as the predominant pair after an initial phase during which Cournot, then 0.3 predominate. However with an intermediate level of noise as in fig. 18, the 0.3 becomes predominant and average profits tend to 0.12. With high noise (fig. 17) however we observe a cycle with 0.3 and the Cournot location with a combined share of about 0.4 and the rest being a mixture of other symmetric and non symmetric locations. Because of the cycle average profits do not converge, but also cycle a little below the Cournot output.

5 Conclusions and Extensions

In this paper we have adopted a simulation methodology to explore how robust the theoretical results of Dixon (1996) are concerning the learning model with a known average, as well as exploring how the model behaves dynamically. Each of the three switching models has interesting and distinct features. The BR dynamic was very different to the other two, in the sense that the switching probabilities are non-zero only for best responses. This can give rise to market outcomes that are very different to the collusive outcome, and in particular with small amounts of noise the Cournot pair has a large population share. However, the BR is a very rigid rule that requires a firm to “switch” to its best responses even when it is currently choosing a the same best response. Experimental behaviour is very limited. The IM and the RS switching regimes give rise to the JPM outcome when there is little noise, and allow for the real possibility of experimentation across a wide range (possibly all) of the available strategies. The interesting difference between IM and the other two switching rules is that it is path dependent, and the switching probabilities depend on the current population proportions of strategies.

Clearly, our simulations can be generalised in several ways. We will list a few of the most important generalisations. First, the noise parameters δ and β can be made to depend on time and the aspiration level. Early on, people

tend to experiment more, and later on less: this can be captured by having δ decreasing over time and possibly tending to zero, and β starting off small and increasing to some upper limit strictly below unity. The probability of experimentation could also be made to depend on the distance of current profits from the aspiration level. Thus δ and β could be combined into a single variable, the experimentation probability ϵ , which is a decreasing function of the difference between current profits and the aspiration level. The extreme case in Dixon (1996) is $\epsilon = 0$ when profits are at least at the aspiration level, and 1 when they are below: the function could also be time dependent. Secondly, the dynamics of the aspiration level could be more complicated (for example a partial adjustment model). These and further extensions await further research.

A Appendix

All the simulations were run on a Sun Sparcstation 5 using Gauss ver. 3.2.18 for Solaris 2.4. In a few cases the output of Gauss was post-processed by Mathematica ver 2.21 for Windows.

The following table summarises the results of 42 of the many simulations we run. The first column of the table indicates the simulation. The second column specifies the learning rule adopted (the simulations marked by † were run on the large specification of the set of permissible outputs). Columns 3 to 5 contain the values of γ , β and δ noise (where the values in columns 4 and 5, when different from 0, are the exponents of $1.0e^{-x}$); columns 6 to 9 the shares of the population at the four reference pairs, and finally columns 10 the number of iterations necessary to stop the the evolutionary process.

Table 5: Details of the simulations

(1) N.	(2) Rule	(3) γ	(4) δ	(5) β	(6) 0.2	(7) JPM	(8) 0.3	(9) Cournot	(10) Iter.
1	RS †	0	0	0	0.0001	0.9862	0.0000	0.0000	4633
2	RS	0	4	4	0.0002	0.9750	0.0003	0.0003	1019
3	RS	0	3	3	0.2376	0.3000	0.3890	0.0009	309
4	RS	0	2	2	0.0773	0.1150	0.1168	0.1850	1548
5	RS	0	1	1	0.0197	0.0259	0.0271	0.0316	34
6	RS	0	4	0	0.0001	0.9787	0.0002	0.0002	1569
7	RS	0	3	0	0.2340	0.2991	0.3951	0.0009	968
8	RS	0	2	0	0.0785	0.1167	0.1175	0.1849	637
9	RS	0	1	0	0.0222	0.0279	0.0279	0.0301	60
10	IM †	0	0	0	0.0000	0.9986	0.0000	0.0000	1778
11	IM	0	4	4	0.0000	0.9931	0.0000	0.0000	311
12	IM	0	3	3	0.0000	0.9931	0.0000	0.0000	316
13	IM	0	2	2	0.0000	0.9929	0.0000	0.0000	336
14	IM	0	1	1	0.0000	0.0000	0.9902	0.0000	414
15	IM	0	4	0	0.0000	0.9931	0.0000	0.0000	311
16	IM	0	3	0	0.0000	0.9930	0.0000	0.0000	312
17	IM	0	2	0	0.0000	0.9930	0.0000	0.0000	330
18	IM	0	1	0	0.0000	0.0000	0.9918	0.0000	347
19	BR †	0	0	0	0.0000	0.0173	0.0260	0.2338	8
20	BR	0	4	4	0.0402	0.0704	0.0704	0.2817	4102
21	BR	0	3	3	0.0073	0.0128	0.0129	0.7662	2111
22	BR	0	2	2	0.0000	0.0001	0.0001	0.9973	752
23	BR	0	1	1	0.0000	0.0000	0.0000	1.0000	148

Table 5: Details of the simulations

(1) N.	(2) Rule	(3) γ	(4) δ	(5) β	(6) 0.2	(7) JPM	(8) 0.3	(9) Cournot	(10) Iter.
24	BR	0	4	0	0.0402	0.0704	0.0704	0.2818	4102
25	BR	0	3	0	0.0005	0.0009	0.0009	0.9700	4730
26	BR	0	2	0	0.0000	0.0001	0.0001	0.9973	751
27	BR	0	1	0	0.0000	0.0000	0.0000	0.9998	100
28	IM/RS \dagger	0.25	0	0	0.0000	1.0000	0.0000	0.0000	326
29	IM/RS \dagger	0.5	0	0	0.0000	1.0000	0.0000	0.0000	209
30	IM/RS \dagger	0.75	0	0	0.0000	1.0000	0.0000	0.0000	231
31	IM/RS	0.5	4	4	0.0000	0.9995	0.0000	0.0000	105
32	IM/RS	0.5	3	3	0.0000	0.9956	0.0000	0.0000	107
33	IM/RS	0.5	2	2	0.0036	0.0048	0.9578	0.0004	340
34	IM/RS	0.5	1	1	0.0041	0.0104	0.1611	0.1550	196
35	IM/RS	0.5	4	0	0.0000	0.9995	0.0000	0.0000	105
36	IM/RS	0.5	3	0	0.0000	0.9956	0.0000	0.0000	107
37	IM/RS	0.5	2	0	0.0036	0.0048	0.9584	0.0003	339
38	IM/RS	0.5	1	0	0.0044	0.0100	0.1399	0.1986	760
39	BR/RS	0.01	0	0	0.0000	0.9948	0.0000	0.0039	9343
40	BR/RS	0.1	0	0	0.0000	0.9955	0.0000	0.0026	1144
41	BR/RS	0.5	0	0	0.0000	0.9981	0.0000	0.0003	516

References

- Banerjee, A. (1992). A simple model of herd behaviour. *Quarterly Journal of Economics*, 110:797–817.
- Bendor, J. and Mookherjee, D. (1994). Aspirations, adaptive learning and cooperation in repeated games. *CENTER discussion paper 9442*.
- Borgers, T. and Sarin, R. (1994). Learning through reinforcement and the replicator dynamics. *Mimeo UCL*.
- Canning, D. (1992). Average behaviour in learning models. *Journal of Economic Theory*, 57:442–472.
- Dixon, H. (1996). Keeping up with the joneses: Aspirations and learning with a known average. *Discussion papers in Economics, University of York*.
- Ellison, G. (Sep 1993). Learning, local interaction, and coordination. *Econometrica*, 61(5):1047.
- Kandarikar, R., Mookherjee, D., Ray, D., and Vega-Redondo, F. (1995). Evolving aspirations and cooperation. *Mimeo, University of Alicante*.
- Kornai, J. (1971). *Anti-equilibrium*. North-Holland / Elsevier.
- Lewin, K. (1936). *Principles of Topological Psychology*. McGraw Hill, New York.
- Lupi, P. (1997). Comparing different learning mechanisms in symmetric bimatrix games. *Mimeo, University of York*.
- Marimon, R. (1995). Learning from learning. *Mimeo EUI Florence*.
- Marimon, R. and grattan, E. (1995). On adaptive learning in strategic games. In Kirman and Salmon, editors, *Learning and Rationality in Economics*, chapter 3. Blackwells.
- Palomino, F. and Vega-Redondo, F. (1996). Convergence of aspirations and (partial) cooperation in the prisoner’s dilemma. *Mimeo University of Alicante*.
- Schlag, K. (1996a). Which one should i imitate? *Mimeo University of Bonn*.
- Schlag, K. (1996b). Why imitate, and if so, how? *Discussion paper*.

- Selten, R. (1991). Evolution, learning and economic behaviour. *Games and Economic Behavior*, 3:3–24.
- Siegel, S. (1957). Level of aspiration and decision making. *Psychological review*, 64:253–262.
- Simon, H. (1947). *Administrative Behaviour: a Study of Decision-making Processes in Administrative Organizations*. MacMillan.
- Simon, H. (1981). *The Science of the Artificial*. M.I.T. press.
- Weibull, J. W. (1995). *Evolutionary Game Theory*. MIT Press.

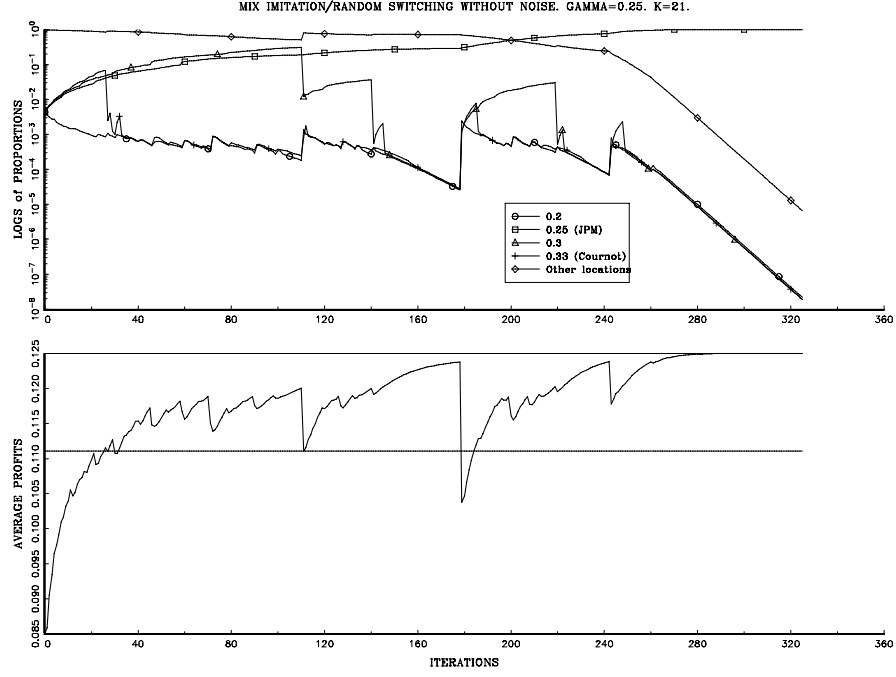


Figure 14: Mixed switching rule: IM and RS ($\gamma = 0.25$).

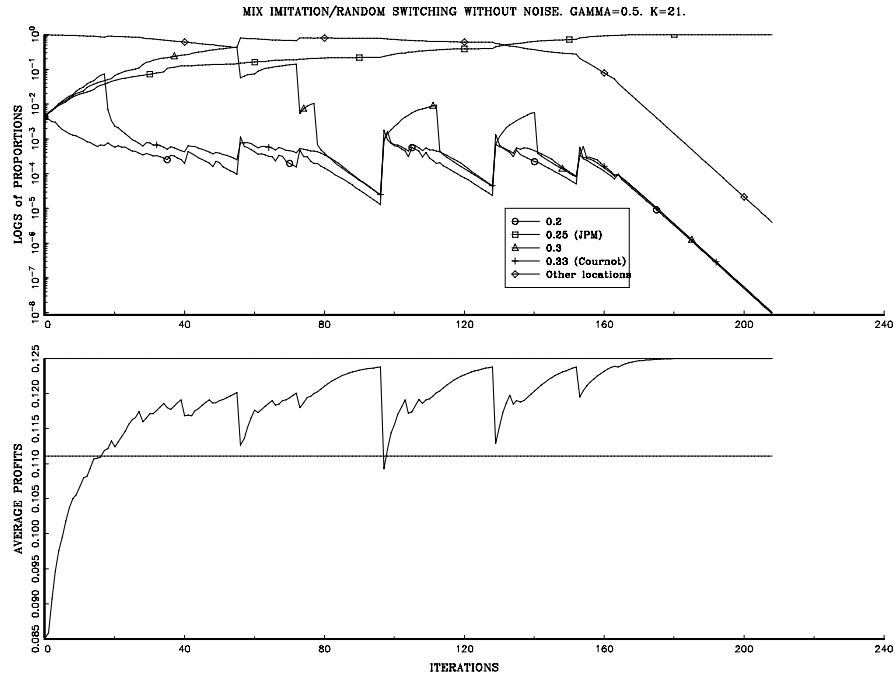


Figure 15: Mixed switching rule: IM and RS ($\gamma = 0.5$).

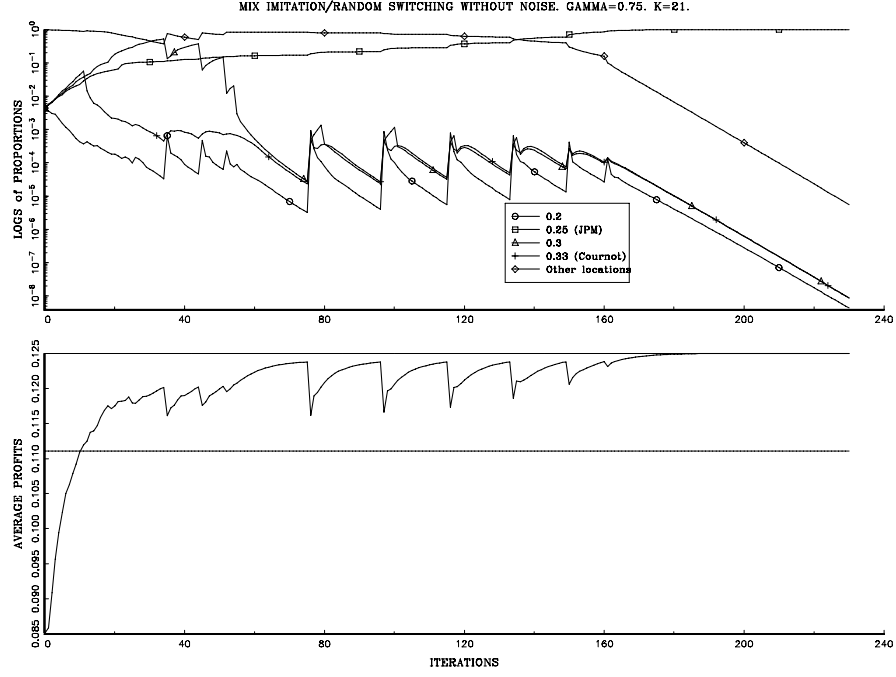


Figure 16: Mixed switching rule: IM and RS ($\gamma = 0.75$).

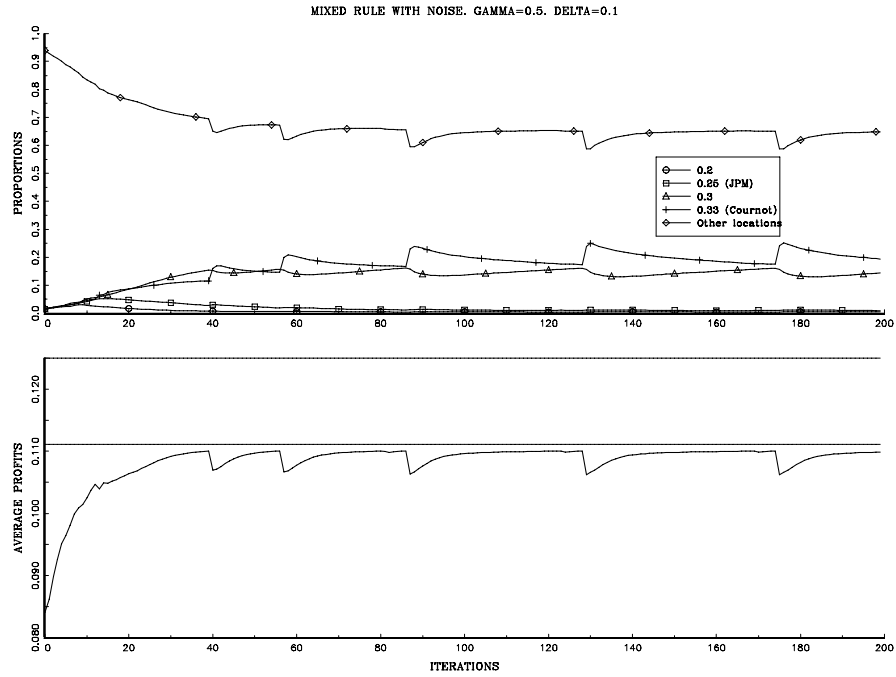


Figure 17: Mixed rule with noise: IM and RS ($\gamma = 0.5$, $\delta = 0.1$).

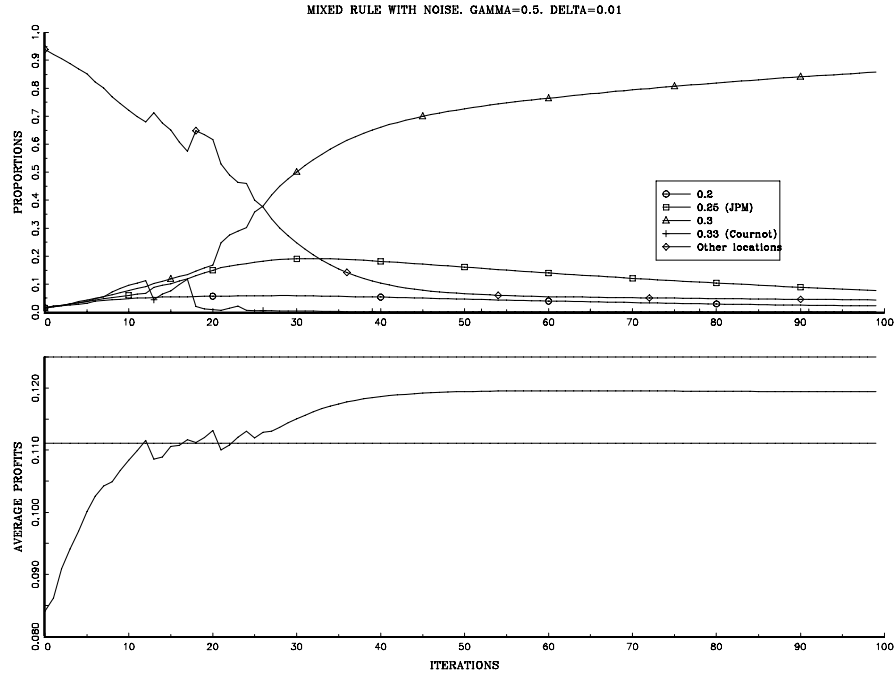


Figure 18: Mixed rule with noise: IM and RS ($\gamma = 0.5$, $\delta = 0.01$).

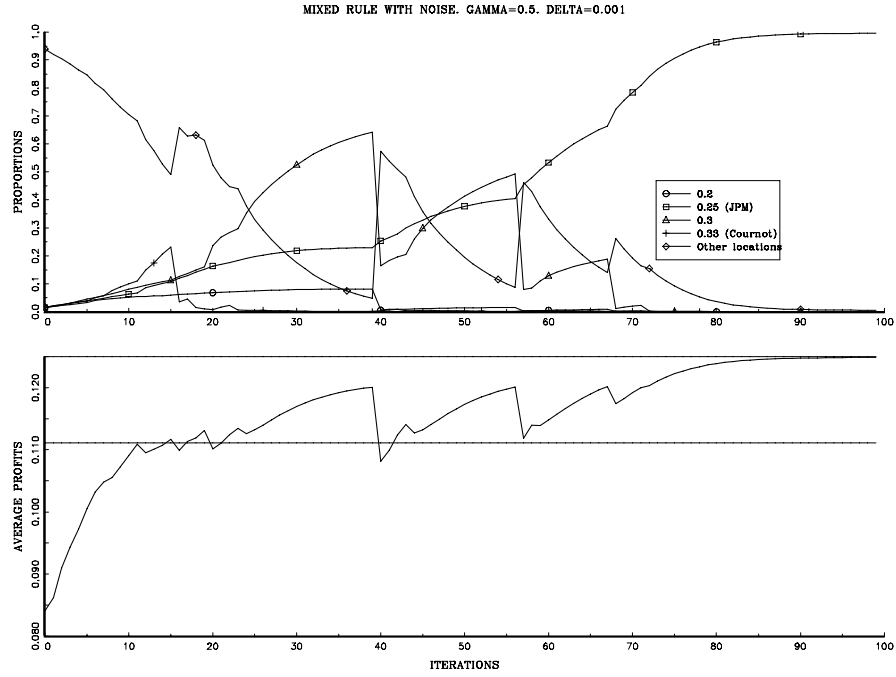


Figure 19: Mixed rule with noise: IM and RS ($\gamma = 0.5$, $\delta = 0.001$).